

Topological study and Lyapunov exponent of a secure steganographic scheme

Nicolas Friot, Christophe Guyeux, and Jacques M. Bahi

Computer science laboratory DISC
FEMTO-ST Institute, UMR 6174 CNRS
University of Franche-Comté, Besançon, France

{nicolas.friot, christophe.guyeux, jacques.bahi}@femto-st.fr

March 8, 2013

Abstract

CIS_2 is a steganographic scheme proposed in the information hiding literature, belonging into the small category of algorithms being both stego and topologically secure. Due to its stego-security, this scheme is able to face attacks that take place into the “watermark only attack” framework. Its topological security reinforce its capability to face attacks in other frameworks as “known message attack” or “known original attack”, in the Simmons’ prisoner problem. In this research work, the study of topological properties of CIS_2 is enlarged by describing this scheme as iterations over the real line, and investigating other security properties of topological nature as the Lyapunov exponent. Results show that this scheme is able to withdraw a malicious attacker in the “estimated original attack” context too.

Keywords: Information Hiding, Steganography, Security, Topology.

1 Introduction

The first fundamental work in information hiding security was realized by Cachin in the early ‘00s, in the context of steganography [6]: attempts of an attacker to make the distinction between an innocent image and a stego-content was rewritten in this article as a hypothesis testing problem. The basic properties of a stegosystem are defined by Cachin using the notions of entropy, mutual information, and relative entropy. At the same time, Mittelholzer has proposed the first theoretical framework for analyzing the security in the second category of algorithms studied by the information hiding community, namely the digital watermarking [17]. These efforts to bring a theoretical framework for security in steganography and watermarking have been followed up

by Kalker, who tries to clarify the concepts (robustness vs. security), and the classifications of watermarking attacks [14]. This work has been deepened by Furon *et al.*, who have translated Kerckhoffs' principle (Alice and Bob shall only rely on some previously shared secret for privacy), from cryptography to data hiding [12]. They used Diffie and Hellman methodology, and Shannon's cryptographic framework [20], to classify the watermarking attacks into categories, according to the type of information Eve has access to [18], namely: Watermarked Only Attack (WOA), Known Message Attack (KMA), Known Original Attack (KOA), Constant-Message Attack (CMA), and Estimated Original Attacks (EOA).

Levels of security have been recently defined in these setups. The highest level of security in WOA is called stego-security [7], whereas topological security tends to improve the ability to withstand attacks in KMA, KOA, and CMA setups [11]. In order to enlarge our knowledge of the level of security of a steganographic scheme, the quantity of disorder generated by the chaos of its topological security can be measured evaluating the well-known Lyapunov exponent [16, 15, 1]. The evaluation of this exponent allow to characterize the ability of the scheme to face an attacker in the context of an EOA.

The first contribution of this article consists in some relation established between how fine is a topology and the chaotic behavior of a dynamical system described with this topology. The second contribution is the security study of a previously released stego and topologically secure steganographic scheme called CIS_2 , on a new topological space, namely the real line numbers \mathbb{R} . On this new space, the topological security of CIS_2 is firstly evaluated, and its Lyapunov exponent is then computed, in order to quantify its level of disorder. Incidentally, this computation allows to measure the resistance of the CIS_2 scheme against a category of attacks called Estimated Original Attack. This study follows a same canvas than a previous work dealing with digital watermarking, but it is conducted here for a steganographic scheme. To achieve this work, a new semi-conjugacy model must be written for the scheme CIS_2 , which is then established and proven here.

This document is organized as follows. In Section 2, some basic notations and terminologies are given and definitions taken from the mathematical topology are recalled. Notions and firsts results concerning the mathematical theory of chaos are introduced in the next section. Then, in Section 4, security notions and classes of attacks under consideration in the information hiding community are recalled. In Section 5, the steganographic scheme studied in this document is presented, and the formalization allowing its topological security evaluation is given in the next section. This model is then used in Section 7 to design a new semi-conjugacy allowing its security study on a new space (Sect. 8), and its Lyapunov exponent is finally evaluated in Section 9. This paper ends by a conclusion section where our contribution is summarized and intended future researches are given.

2 Basic Reminders

In what follows, \mathbb{B} denotes the Boolean set $\{0, 1\}$, S^n stands for the n^{th} term of a sequence S , V_i is for the i^{th} component of a vector V , and $\llbracket 0; N \rrbracket$ is the integer interval

$\{0, 1, \dots, N\}$. Furthermore, the following definitions will be used in this document.

Definition 1 The *discrete Boolean metric* is the application $\delta : \mathbb{B} \rightarrow \mathbb{B}$ defined by $\delta(x, y) = 0 \Leftrightarrow x = y$.

Definition 2 The *vectorial negation* is the function $f_0 : \mathbb{B}^N \rightarrow \mathbb{B}^N$ defined by $f_0((b_0, \dots, b_{N-1})) = (\overline{b_0}, \dots, \overline{b_{N-1}})$, where \overline{x} is the negation of the Boolean x .

Definition 3 A *strategy adapter* is a sequence which elements belong into $\llbracket 1, k \rrbracket$, where $k \in \mathbb{N}^*$. The set of all strategies having terms in $\llbracket 1, k \rrbracket$ is denoted by \mathbb{S}_k .

Definition 4 For $k \in \mathbb{N}^*$, the *initial function* is the map $i_k : \mathbb{S}_k \rightarrow \llbracket 1, k \rrbracket$ defined by $i_k((S^n)_{n \in \mathbb{N}}) = S^0$.

Definition 5 Let $k \in \mathbb{N}^*$. The *shift function* is the map $\sigma_k : \mathbb{S}_k \rightarrow \mathbb{S}_k$ defined by $\sigma_k((S^n)_{n \in \mathbb{N}}) = (S^{n+1})_{n \in \mathbb{N}}$.

After having introduced these definitions, we will now give some recalls in the field of the mathematical topology [19] to make this document self contained, and because topology is not of frequent use in cryptology.

3 A few Words Concerning the Mathematical Theory of Chaos

3.1 Devaney's Chaotic Dynamical Systems

Some topological definitions and properties taken from the mathematical theory of chaos are now recalled in this section. Let (X, τ) be a topological space and f a continuous function on (X, τ) .

Definition 6 f is said to be *topologically transitive* if, for any pair of open sets $U, V \subset X$, there exists $k > 0$ such that $f^k(U) \cap V \neq \emptyset$.

Definition 7 (X, f) is *regular* if the set of periodic points is dense in X .

It is now possible to introduce the well-established mathematical definition of chaos [8].

Definition 8 A function $f : X \rightarrow X$ is said to be *chaotic* on X if it is regular and topologically transitive.

If the topological space is indeed a metric space (X, d) , then the sensibility of the system under iterations, regarding its initial conditions, can be quantified as follows.

Definition 9 f has *sensitive dependence on initial conditions* if there exists $\delta > 0$ such that, for any $x \in X$ and any neighborhood V of x , there exist $y \in V$ and $n \geq 0$ such that $d(f^n(x), f^n(y)) > \delta$. δ is called the *constant of sensitivity* of f .

This property is implied by both the regularity and transitivity presented above [5]. And so, when f is chaotic, fundamentally different behaviors are possible for the system, and they occur in an unpredictable way.

Let us state now some basic results that surprisingly cannot be found in the literature. To simplify the presentation, some notations must be firstly introduced: X_τ will stand for the topological space (X, τ) , whereas $\mathcal{V}_\tau(x)$ is the set of neighborhoods of x for the topology τ (in unambiguous cases, we will simply use $\mathcal{V}(x)$).

Theorem 1 *Let X be a set, and τ, τ' two topologies on X such that τ' is finer than τ . Let $f : X \rightarrow X$, continue for both τ and τ' .*

If $(X_{\tau'}, f)$ is chaotic in the sense of Devaney, then (X_τ, f) is also chaotic.

Proof 1 *Let ω_1, ω_2 two open sets of τ . Then $\omega_1, \omega_2 \in \tau'$, because τ' is finer than τ . As f is τ' -transitive, then $\exists n \in \mathbb{N}, \omega_1 \cap f^{(n)}(\omega_2) = \emptyset$. As a consequence, f is τ -transitive.*

Let us now establish the regularity of (X_τ, f) , i.e., for all $x \in X$ and all τ -neighborhood V of x , there exists a periodic point for f in V . Let $x \in X$, and $V \in \mathcal{V}_\tau(x)$ a τ -neighborhood of x . By definition of a neighborhood, $\exists \omega \in \tau, x \in \omega \subset V$. However $\tau \subset \tau'$, so $\omega \in \tau'$, and so $V \in \mathcal{V}_{\tau'}(x)$. But $(X_{\tau'}, f)$ is regular, So there exists a periodic point for f in V , and the regularity of (X_τ, f) is proven.

Let us finally recall another topological quantitative property of chaos:

Definition 10 A function f is said to be *expansive* if $\exists \epsilon > 0, \forall x \neq y, \exists n \in \mathbb{N}, d(f^n(x), f^n(y)) \geq \epsilon$.

Sometimes, instead of trying to prove properties directly on the system itself, it is preferable to reduce the initial problem to another one whose characteristics are known or appear more accessible. Such a reduction tool is called, in the mathematical theory of chaos, the semi-conjugacy.

3.2 The topological semi-conjugacy

Definition 1 *The discrete dynamical system (X, f) is topologically semi-conjugate to the system (\mathcal{Y}, g) if it exists a function $\phi : X \rightarrow \mathcal{Y}$, both continuous and onto, such that: $\phi \circ f = g \circ \phi$, that is, which makes commutative the following diagram [10]:*

$$\begin{array}{ccc} X & \xrightarrow{f} & X \\ \phi \downarrow & & \downarrow \phi \\ \mathcal{Y} & \xrightarrow{g} & \mathcal{Y} \end{array}$$

In this case, the system (\mathcal{Y}, g) is called a factor of the system (X, f) .

Several dynamic behaviors are inherited by systems factors [10]. They are summarized in the following proposition:

Proposition 1 *Let (\mathcal{Y}, g) a factor of the system (X, f) . Then:*

1. $p \in \text{Per}_k(f) \implies \varphi(p) \in \text{Per}_j(g)$, where $j \leq k$, where $\text{Per}_n(h)$ stands for the set of points of period n for the iteration function h .
2. (X, f) regular $\implies (\mathcal{Y}, g)$ regular;
3. (X, f) transitive $\implies (\mathcal{Y}, g)$ transitive.

So if (X, f) is chaotic in the sense of Devaney, then (\mathcal{Y}, g) is chaotic too.

3.3 The Lyapunov Exponent

Some dynamical systems are very sensitive to small changes in their initial conditions, which is illustrated by both the constants of sensitivity to initial conditions and of expansivity introduced respectively in Definitions 9 and 10. However, these variations can quickly take enormous proportions, grow exponentially, and none of these constants can illustrate that. Alexander Lyapunov has examined this phenomenon and introduced an exponent that measures the rate at which these small variations can grow.

Definition 11 Given $f : \mathbb{R} \longrightarrow \mathbb{R}$, the *Lyapunov exponent* of the system composed by $x^0 \in \mathbb{R}$ and $x^{n+1} = f(x^n)$ is defined by:

$$\lambda(x_0) = \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{i=1}^n \ln \left| f'(x^{i-1}) \right|.$$

Consider a dynamical system with an infinitesimal error on the initial condition x_0 . When the Lyapunov exponent is positive, this error will increase exponentially (situation of chaos), whereas it will decrease if $\lambda(x_0) \leq 0$.

Let us now recall the information hiding security framework developed this last decade.

4 Information Hiding Security

In the prisoner problem of Simmons [21], Alice and Bob are in jail, and they want to, possibly, devise an escape plan by exchanging hidden messages in innocent-looking cover contents. These messages are to be conveyed to one another by a common warden, Eve, who over-drops all contents and can choose to interrupt the communication if they appear to be stego-contents.

In the steganography framework, in the context of the Simmons' prisoner problem, attacks have been classified in [7] as follows:

Definition 12 WOA: A *Watermark-Only Attack* occurs when an attacker has only access to several watermarked contents.

KMA: A *Known-Message Attack* occurs when an attacker has access to several pairs of watermarked contents and corresponding hidden messages.

KOA: A *Known-Original Attack* is when an attacker has access to several pairs of watermarked contents and their corresponding original versions.

CMA: A *Constant-Message Attack* occurs when the attacker observes several watermarked contents and only knows that the unknown hidden message is the same in all contents.

EOA: An *Estimated Original Attacks* occurs when the attacker has access to an estimation of the original host signal, with possibly some estimation errors.

In the framework of WOA, the stego-security [7] is relevant to evaluate the security of information hiding processes. It is the highest security level in WOA setup. To recall it, we need the following notations: \mathbb{K} is the set of embedding keys, $p(X)$ is the probabilistic model of N_0 initial host contents, $p(Y|K_1)$ is the probabilistic model of N_0 watermarked contents. Furthermore, it is supposed in this context that each host content has been watermarked with the same secret key K_1 and the same embedding function e . It is now possible to define the notion of stego-security:

Definition 13 (Stego-Security) The embedding function e is *stego-secure* if and only if:

$$\forall K_1 \in \mathbb{K}, p(Y|K_1) = p(X).$$

In the other frameworks, KOA, KMA, and CMA, the topological security should be investigated [11]. In this article, we focus more specifically on this last point, which is recalled below.

To check whether an information hiding scheme S is topologically secure or not, S must be written as an iterate process $x^{n+1} = f(x^n)$ on a metric space (X, d) . This formulation is always possible [4]. So,

Definition 14 (Topological Security) An information hiding scheme S is said to be topologically secure on (X, d) if its iterative process has a chaotic behavior according to Devaney.

In the approach presented by Guyeux *et al.*, a data hiding scheme is secure if it is unpredictable. Its iterative process must satisfy the Devaney's chaos property and its level of topological security increases with the number of chaotic properties satisfied by it.

This new concept of security for data hiding schemes has been proposed in [4] as a complementary approach to the existing framework. It contributes to the reinforcement of confidence into existing secure data hiding schemes. Additionally, the study of security in KMA, KOA, and CMA setups is realizable in this context. Finally, this framework can replace stego-security in situations that are not encompassed by it. In particular, this framework is more relevant to give evaluation of data hiding schemes claimed as chaotic.

In the framework of the last attack EOA, the evaluation of the Lyapunov exponent, which is the subject of this research work, is relevant to quantify the level of security of steganographic processes proven to be topologically secure. Indeed, the Lyapunov exponent participates to the measurement of this topological security. In an EAO setup, the attacker has just access to estimation of the original content. With just this knowledge, he should not be in measure to recover any information about the secret message

or the secret key. The topological security with the two notions of sensibility and expansivity introduced in Definitions 9 and 10 are relevant to face the attacker in this context. However, these two topological properties give no precise quantification. It is the major interest of the Lyapunov exponent.

5 The Steganographic Scheme CIS_2

To explain how to use chaotic iterations for information hiding, we must firstly define the significance of a given coefficient, and the notion of most and least significant coefficients (MSCs and LSCs).

We first notice that terms of the original content x that may be replaced by terms issued from the secret message y are less important than others: they could be changed without be perceived as such. For a given host content x , MSCs are then ranks of x that describe the relevant part of the image, whereas LSCs translate its less significant parts. These two definitions are illustrated on Figure 1, where the LSCs correspond to the last three bits of each pixel.

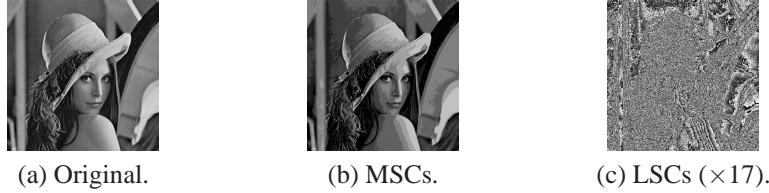


Figure 1: Most and least significant coefficients of Lena.

The steganographic scheme CIS_2 that generalize the watermarking scheme based on chaotic iterations as described in [1] can now be recalled. In this part the following notations will be used: $x^0 \in \mathbb{B}^N$ is the N LSCs of a given cover media C , $m^0 \in \mathbb{B}^P$ is the secret message to embed into x^0 , $S_p \in \mathbb{S}_N$ is the *place strategy*, $S_c \in \mathbb{S}_P$ is the *choice strategy*, and lastly, $S_m \in \mathbb{S}_P$ is a the *mixing strategy*.

The steganographic scheme is defined by $\forall (n, i, j) \in \mathbb{N}^* \times \llbracket 0; N-1 \rrbracket \times \llbracket 0; P-1 \rrbracket$:

$$x_i^n = \begin{cases} x_i^{n-1} & \text{if } S_p^n \neq i \\ m_{S_c^n}^n & \text{if } S_p^n = i, \end{cases} \text{ and } m_j^n = \begin{cases} m_j^{n-1} & \text{if } S_m^n \neq j \\ \overline{m_j^{n-1}} & \text{if } S_m^n = j. \end{cases}$$

where $\overline{m_j^{n-1}}$ is the Boolean negation of m_j^{n-1} .

The new LSCs of the stego-content are the Boolean vector $y = x^P \in \mathbb{B}^N$.

6 Topological Model of CIS_2 and Security on \mathcal{X}_2

In this section is recalled the topology used in order to model the steganographic scheme CIS_2 by a discrete dynamical system in a topological space [8].

Let

$$\begin{aligned} \mathcal{F} : \quad & \llbracket 0; N-1 \rrbracket \times \mathbb{B}^N \times \llbracket 0; P-1 \rrbracket \times \mathbb{B}^P \longrightarrow \mathbb{B}^N \\ & (k, x, \lambda, m) \longmapsto \left(\delta(k, j) \cdot x_j + \overline{\delta(k, j)} \cdot m_{\lambda_j} \right)_{j \in \llbracket 0; N-1 \rrbracket} \end{aligned}$$

where $+$ and \cdot are the Boolean addition and product.

Consider the phase space \mathcal{X}_2 defined as follow: $\mathcal{X}_2 = \mathbb{S}_N \times \mathbb{B}^N \times \mathbb{S}_P \times \mathbb{B}^P \times \mathbb{S}_P$, where \mathbb{S}_N and \mathbb{S}_P are the sets introduced in Section 2.

The map $\mathcal{G}_{f_0} : \mathcal{X}_2 \longrightarrow \mathcal{X}_2$ is defined by: $\mathcal{G}_{f_0}(S_p, x, S_c, m, S_m) = (\sigma_N(S_p), \mathcal{F}(i_N(S_p), x, i_P(S_c), m), \sigma_P(S_c), \mathcal{G}_{f_0}(m, S_m), \sigma_P(S_m))$. Then CIS_2 can be described by the iterations of the following discrete dynamical system:

$$X^0 \in \mathcal{X}_2 \text{ and } X^{k+1} = \mathcal{G}_{f_0}(X^k).$$

By comparing \mathcal{X}_2 and \mathcal{X}_1 , it has been proven in [11] that:

Proposition 1 \mathcal{X}_2 has, at least, the cardinality of the continuum.

A new distance has been defined on \mathcal{X}_2 as follow: $\forall X, \check{X} \in \mathcal{X}_2$, if $X = (S_p, x, S_c, m, S_m)$ and $\check{X} = (\check{S}_p, \check{x}, \check{S}_c, \check{m}, \check{S}_m)$, then: $d_2(X, \check{X}) = d_{\mathbb{B}^N}(x, \check{x}) + d_{\mathbb{B}^P}(m, \check{m}) + d_{\mathbb{S}_N}(S_p, \check{S}_p) + d_{\mathbb{S}_P}(S_c, \check{S}_c) + d_{\mathbb{S}_P}(S_m, \check{S}_m)$, where: $d_{\mathbb{B}^N}(E, \check{E}) = \sum_{k=0}^{N-1} \delta(E_k, \check{E}_k) \in \llbracket 0; N \rrbracket$ and $d_{\mathbb{S}_N}(S, \check{S}) =$

$$\frac{9}{N} \sum_{k=1}^{\infty} \frac{|S^k - \check{S}^k|}{10^k} \in [0; 1] \text{ are respectively distances on } \mathbb{B}^N \text{ and } \mathbb{S}_N \text{ } (\forall N \in \mathbb{N}^*).$$

To demonstrate that CIS_2 is another example of topological chaos in the sense of Devaney, it has been firstly established in [11] that,

Proposition 2 \mathcal{G}_{f_0} is continuous on (\mathcal{X}_2, d_2) .

Then it has been proven that $(\mathcal{X}_2, \mathcal{G}_{f_0})$ is topologically transitive, regular, and has sensitive dependence on initial conditions. Then we have the result:

Theorem 1 \mathcal{G}_{f_0} is a chaotic map on (\mathcal{X}_2, d_2) in the sense of Devaney, and consequently the scheme CIS_2 is topologically secure.

An other theorem about the security of CIS_2 can be given here. This result has been established in [11].

Theorem 2 CIS_2 is stego-secure.

7 A Topological Semi-conjugacy between \mathcal{X}_2 and \mathbb{R}

In this section, by using a topological semi-conjugacy, we show that CIS_2 modeled by \mathcal{G}_{f_0} on \mathcal{X} can be described as iterations on a real interval. As our researches are inspired by the work of the authors of [1], the proofs detailed in this document will follow a same canvas. To do so, some notations and terminologies must be introduced.

Let $\mathcal{X}_{(N;P)} = \mathbb{S}_N \times \mathbb{B}^N \times \mathbb{S}_P \times \mathbb{B}^P \times \mathbb{S}_P$. In what follows and for easy understanding, we will assume that $N = 3$ and $P = 2$. Indeed, N has obviously to be greater than P . So $N + P = 5$ and $NP^2 = 12$. However, an equivalent formulation of the following can be easily obtained by replacing the bases 5 and 12 by any base $(N + P)$ and (NP^2) .

Definition 15 The function $\psi : \llbracket 1, N \rrbracket \times \llbracket 1, P \rrbracket \times \llbracket 1, P \rrbracket \rightarrow \llbracket 0, NP^2 - 1 \rrbracket$ is defined by:

$$\psi(S_p^i, S_c^i, S_m^i) = (S_p^i - 1)P^2 + (S_c^i - 1)P + (S_m^i - 1).$$

This function aims to convert a triple strategy value in a simple strategy expressed in another base. ψ is a bijective function. The reverse operation will be denoted by ψ^{-1} . The three projections of ψ^{-1} are so defined by: $\psi_1^{-1}(\psi(S_p^i, S_c^i, S_m^i)) = S_p^i$,

$$\psi_2^{-1}(\psi(S_p^i, S_c^i, S_m^i)) = S_c^i, \text{ and } \psi_3^{-1}(\psi(S_p^i, S_c^i, S_m^i)) = S_m^i.$$

| Base N = 3 | Base P = 2 | Base P = 2 | Base NP ² = 12 | Base N = 3 | Base P = 2 | Base P = 2 | Base NP ² = 12 |
|---------------|---------------|---------------|------------------------------|---------------|---------------|---------------|------------------------------|
| S_p^i | S_c^i | S_m^i | $\psi(S_p^i, S_c^i, S_m^i)$ | S_p^i | S_c^i | S_m^i | $\psi(S_p^i, S_c^i, S_m^i)$ |
| 1 | 1 | 1 | 0 | 2 | 2 | 1 | 6 |
| 1 | 1 | 2 | 1 | 2 | 2 | 2 | 7 |
| 1 | 2 | 1 | 2 | 3 | 1 | 1 | 8 |
| 1 | 2 | 2 | 3 | 3 | 1 | 2 | 9 |
| 2 | 1 | 1 | 4 | 3 | 2 | 1 | 10 |
| 2 | 1 | 2 | 5 | 3 | 2 | 2 | 11 |

Table 1: Illustration of the function ψ (see Definition 15).

Definition 16 The function $\phi : \mathcal{X}_{(3;2)} \rightarrow [0, 2^5[$ is defined by: $\phi : \mathcal{X}_{(3;2)} = \mathbb{S}_3 \times \mathbb{B}^3 \times$

$\mathbb{S}_2 \times \mathbb{B}^2 \times \mathbb{S}_2 \longrightarrow [0, 2^5[$ where $(S_p, E, S_c, M, S_m) = ((S_p^0, S_p^1, \dots); (E_0, E_1,$

$E_3); (S_c^0, S_c^1, \dots); (M_0, M_1); (S_m^0, S_m^1, \dots))$, and $\phi(S_p, E, S_c, M, S_m)$ is the real number whose integral part e is $\sum_{k=0}^2 2^{4-k} E_k + \sum_{k=3}^4 2^{4-k} M_{k-3}$, that is, the binary digits of e are $E_0 E_1 E_2 M_0 M_1$, and whose decimal part s is equal to: $s = 0, \psi(S_p^0, S_c^0, S_m^0) \psi(S_p^1, S_c^1, S_m^1) \psi(S_p^2, S_c^2, S_m^2) \dots = \sum_{k=1}^{+\infty} 12^{-k} S^k$. s is thus expressed in base 12.

ϕ realizes the association between a point of $\mathcal{X}_{(3;2)}$ and a real number into $[0, 2^5[$. We must now translate the steganographic process CIS_2 represented by G_{f_0} iterations on this real interval. To do so, two intermediate functions over $[0, 2^5[$ denoted by e and s must be introduced:

Definition 17 Let $x \in [0, 2^5[$ and e_0, \dots, e_4 the binary digits of the integral part of x : $\lfloor x \rfloor = \sum_{k=0}^4 2^{4-k} e_k$, $(s^k)_{k \in \mathbb{N}}$ the digits of x , expressed in base 12, where the chosen decimal decomposition of x is the one that does not have an infinite number of 11: $x = \lfloor x \rfloor + \sum_{k=0}^{+\infty} s^k 12^{-k-1}$. e and s are thus defined as follows:

$$e : \begin{array}{ccc} [0, 2^5[& \longrightarrow & \mathbb{B}^3 \times \mathbb{B}^2 \\ x & \longmapsto & ((e_0, e_1, e_2); (e_3, e_4)) \end{array} \quad \text{and} \quad s : \begin{array}{ccc} [0, 2^5[& \longrightarrow & \llbracket 0, 11 \rrbracket^{\mathbb{N}} \\ x & \longmapsto & (s^k)_{k \in \mathbb{N}} \end{array}$$

We are now able to define the function g , whose goal is to translate the steganographic process CIS_2 represented by \mathcal{G}_{f_0} on an interval of \mathbb{R} .

Definition 18 $g : [0, 2^5[\longrightarrow [0, 2^5[$ is such that $g(x)$ is the real number of $[0, 2^5[$ defined bellow: its integral part has a binary decomposition equal to e'_0, \dots, e'_4 , with $\forall i \in \llbracket 0, 2 \rrbracket$ and $\forall j \in \llbracket 3, 4 \rrbracket$:

$$e'_i = \begin{cases} e(x)_i & \text{if } i \neq \psi_1^{-1}(s^0) \\ e(x)_{2+\psi_2^{-1}(s^0)} & \text{if } i = \psi_1^{-1}(s^0) \end{cases} \quad e'_j = \begin{cases} e(x)_j & \text{if } j \neq \psi_3^{-1}(s^0) \\ e(x)_j + 1 \pmod{2} & \text{if } j = \psi_3^{-1}(s^0) \end{cases},$$

whose decimal part is $s(x)^1, s(x)^2, \dots$

In other words, if $x = \sum_{k=0}^4 2^{4-k} e_k + \sum_{k=0}^{+\infty} s^k 12^{-k-1}$, then:

$$g(x) = \sum_{k=0}^2 2^{4-k} \left[e_k \left(\delta(k, \psi_1^{-1}(s^0)) + 1 \pmod{2} \right) + e_{2+\psi_2^{-1}(s^0)} \left(\delta(k, \psi_1^{-1}(s^0)) \right) \right] \\ + \sum_{k=3}^4 2^{4-k} (e_k + \delta(k, \psi_3^{-1}(s^0)) \pmod{2}) + \sum_{k=0}^{+\infty} s^{k+1} 12^{-k-1},$$

where δ is the discrete Boolean metric introduced in Definition 1.

Numerous metrics can be defined on the set $[0, 2^5[$, the most usual one being the Euclidian distance $\Delta(x, y) = |y - x|^2$. This Euclidian distance does not reproduce exactly the notion of proximity induced by our first distance d_2 on \mathcal{X}_2 . Indeed d_2 is richer than Δ . This is the reason why we have to introduce the following metric:

Definition 19 Given $x, y \in [0, 2^5[$, D denotes the function from $[0, 2^5[^2$ to \mathbb{R}^+ defined by: $D(x, y) = D_e(e(x), e(y)) + D_s(s(x), s(y))$, where:

$$D_e(e, \check{e}) = \sum_{k=0}^4 \delta(e_k, \check{e}_k), \quad \text{and} \quad D_s(s, \check{s}) = \sum_{k=1}^{\infty} \frac{|s^k - \check{s}^k|}{12^k}.$$

Proposition 3 D is a distance on $[0, 2^5[$.

PROOF The three axioms defining a distance must be checked.

- $D \geq 0$, because everything is positive in its definition. If $D(x, y) = 0$, then $D_e(x, y) = 0$, so the integral parts of x and y are equal (they have the same binary decomposition). Additionally, $D_s(x, y) = 0$, then $\forall k \in \mathbb{N}^*, s(x)^k = s(y)^k$. In other words, x and y have the same k -th decimal digit, $\forall k \in \mathbb{N}^*$. And so $x = y$.
- $D(x, y) = D(y, x)$.

- Finally, the triangular inequality is obtained due to the fact that both δ and $|x - y|$ satisfy it.

The convergence of sequences according to D is not the same than the usual convergence related to the Euclidian metric. For instance, if $x^n \rightarrow x$ according to D , then necessarily the integral part of each x^n is equal to the integral part of x (at least after a given threshold), and the decimal part of x^n corresponds to the one of x “as far as required”. D is richer and more refined than the Euclidian distance, and thus is more precise.

ϕ has been constructed in order to be continuous and onto, so we obtained the following theorem:

Theorem 3 *The steganographic process CIS_2 represented by (G_{f_0}, X_2) is simple iterations on \mathbb{R} , which is illustrated by the semi-conjugacy given below:*

$$\begin{array}{ccc} \left(X_{(3;2)}, d_2 \right) & \xrightarrow{G_{f_0}} & \left(X_{(3;2)}, d_2 \right) \\ \phi \downarrow & & \downarrow \phi \\ \left(\left[0, 2^5 \right[, D \right) & \xrightarrow{g} & \left(\left[0, 2^5 \right[, D \right) \end{array}$$

In other words, X_2 is approximately equal to $\left[0, 2^{N+P} \right[$.

It can be remarked that the function g is a piecewise linear function: it is linear on each interval having the form $\left[\frac{n}{12}, \frac{n+1}{12} \right[$, $n \in \llbracket 0; 2^5 \times 12 \rrbracket$ and its slope is equal to 12. Let us justify these claims:

Proposition 4 *The process CIS_2 represented by g defined on \mathbb{R} has derivatives of all orders on $\left[0, 2^5 \right[$, except on the 385 points in I defined by:*

$$I = \left\{ \frac{n}{12} \mid n \in \llbracket 0; 2^5 \times 12 \rrbracket \right\}.$$

Furthermore, on each interval of the form $\left[\frac{n}{12}, \frac{n+1}{12} \right[$, with $n \in \llbracket 0; 2^5 \times 12 \rrbracket$, g is a linear function, having a slope equal to 12: $\forall x \notin I, g'(x) = 12$.

PROOF Let $I_n = \left[\frac{n}{12}, \frac{n+1}{12} \right[$, with $n \in \llbracket 0; 2^5 \times 12 \rrbracket$. All the points of I_n have the same integral part e and the same decimal part s^0 : on the set I_n , functions $e(x)$ and $x \mapsto s(x)^0$ of Definition 17 only depend on n . So all the images $g(x)$ of these points x :

- Have the same integral part, which is e , except probably the bit number s^0 . In other words, this integer has approximately the same binary decomposition than e , the sole exception being the digit s^0 (this number is then either $e + 2^{12-s^0}$ or $e - 2^{12-s^0}$, depending on the parity of s^0 , i.e., it is equal to $e + (-1)^{s^0} \times 2^{12-s^0}$).
- A shift to the left has been applied to the decimal part y , losing by doing so the common first digit s^0 . In other words, y has been mapped into $12 \times y - s^0$.

To sum up, the action of g on the points of I is as follows: first, make a multiplication by 12, and second, add the same constant to each term, which is $\frac{1}{12} \left(e + (-1)^{s^0} \times 2^{12-s^0} \right) - s^0$.

We are now able to evaluate the Lyapunov exponent of our digital watermarking scheme based on chaotic iterations, which is now described by the iterations on \mathbb{R} of the g function introduced in Definition 18.

8 Topological security of CIS_2 on \mathbb{R}

According to Theorem 1, CIS_2 represented by the function \mathcal{G}_{f_0} on \mathcal{X}_2 is topologically secure, that is to say $(\mathcal{G}_{f_0}, \mathcal{X}_2)$ is chaotic in the sense of Devaney. We can deduce that it is the same for CIS_2 represented by the g function on \mathbb{R} for the order topology. Indeed $(\mathcal{G}_{f_0}, \mathcal{X}_2)$ and $(g, [0, 2^5[_D)$ are semi-conjugate by ϕ as proven in the previous section, so $(g, [0, 2^5[_D)$ is a chaotic system according to Devaney, because the semi-conjugacy preserves the character according to Proposition 1 in Section 3.2. However the topology generated by D is finer than the topology generated by the Euclidean distance Δ – which is the order topology. Finally, according to Theorem 1, we can affirm that the steganographic process CIS_2 represented by g is chaotic in the sense of Devaney for the order topology on \mathbb{R} .

Having these claims in mind, we can formulate the following theorem:

Theorem 2 *The steganographic process CIS_2 represented by g on \mathbb{R} is chaotic in the sense of Devaney, when the usual topology of \mathbb{R} is used (the order topology).*

This result is weaker than Theorem 1, which establish the chaotic property of CIS_2 for a finer topology. It is as if the chaos observed using usual tools like the Euclidean distance is still preserved when considering more powerful tools (higher resolution, *i.e.*, finer topologies).

The result contained in Theorem 2 is however interesting, as it confirms that the followed approach does not lead to weaker properties. Indeed, this study has taken place in a system other than the one usually considered (\mathcal{X}_2 instead of \mathbb{R}), in order to be as closed as possible to the final computer machines. By doing so, we prevent from any loss of chaotic properties when computing the scheme written in mathematical terms. However, it might be feared that the choice of a discrete mathematics approach leads to a disorder of lower quality. In other words, we have achieved to leave a situation of great disorder lost during the computation into machines. However, the cost of such achievement were probably to obtain a disorder of poor quality. Theorem 2 claims exactly the contrary.

9 Evaluation of the Lyapunov Exponent

Let $\mathcal{L} = \left\{ x^0 \in [0, 2^5[\mid \forall n \in \mathbb{N}, x^n \notin I \right\}$, where I is the set of points in the real interval where g is not differentiable (as it is explained in Proposition 4). Then,

Theorem 4 $\forall x^0 \in \mathcal{L}$, the Lyapunov exponent of CIS_2 having x^0 for initial condition is equal to $\lambda(x^0) = \ln(12)$.

PROOF g is piecewise linear, with a slop of 12 ($g'(x) = 12$ where the function g is differentiable). Then $\forall x \in \mathcal{L}$, $\lambda(x) = \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{i=1}^n \ln \left| g'(x^{i-1}) \right|$
 $= \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{i=1}^n \ln |12| = \lim_{n \rightarrow +\infty} \frac{1}{n} n \ln |12| = \ln 12$.

Remark 1 The set of initial conditions for which this exponent is not calculable is countable. This is indeed the initial conditions such that an iteration value will be a number having the form $\frac{n}{12}$, with $n \in \mathbb{N}$. Moreover, for a system having $N + P$ cells (a number of LSCs equal to N and a secret message to embed of width equal to P), we will find, mutatis mutandis, an infinite uncountable set of initial conditions $x^0 \in [0; 2^{N+P}[$ such that $\lambda(x^0) = \ln(NP^2)$.

So, it is possible to make the Lyapunov exponent of the scheme CIS_2 as large as possible, depending on the number of least significant coefficients of the cover media we decide to consider, and on the width of the message to embed. Obviously, a large Lyapunov exponent makes it impossible to achieve the well-known “Estimated Original Attacks” [7].

10 Conclusion and Future Work

To conclude, we have now available a new quantitative evaluation for the steganographic scheme CIS_2 previously proposed, which is based on chaotic iteration: its Lyapunov exponent is equal to $\ln(NP^2)$, where N is the number of least significant coefficients of the cover media and P the width of the secret message to embed. This exponent allows to quantify the amplification of the ignorance on the exact initial condition (the media without watermark) after several iterations of the steganographic process. It illustrates the disorder generated by iterations of the process, reinforcing its chaotic nature. Thanks to its topological security, this scheme is already able to face an attacker in the context of a “Known-Message Attack”, a “Known-Original Attack”, and a “Constant-Message Attack”. In addition, this result implies that it is also able to resist in the context of an “Estimated Original Attacks”.

Using the semi-conjugacy described here, it will be possible in a future work to compare the topological behavior of CIS_2 on \mathcal{X}_2 and \mathbb{R} , and to explore the topological security of the steganography scheme using this new topology. Then, an analogue study of the two other topologically secure schemes cited here will be conducted in order to compare these processes, being thus able to choose the best one according to the type of applications under consideration. Finally, security in steganography context will be investigated too, and topological security will be applied in this framework.

References

- [1] Jacques Bahi, Nicolas Friot, and Christophe Guyeux. Lyapunov exponent evaluation of a digital watermarking scheme proven to be secure. In *IIH-MSP'2012, 8-th Int. Conf. on Intelligent Information Hiding and Multimedia Signal Processing*, pages ***-***, Piraeus-Athens, Greece, July 2012. IEEE Computer Society. To appear.
- [2] Jacques Bahi and Christophe Guyeux. Hash functions using chaotic iterations. *Journal of Algorithms & Computational Technology*, 4(2):167–181, 2010.
- [3] Jacques Bahi and Christophe Guyeux. A new chaos-based watermarking algorithm. In *SECRYPT 2010, International conference on security and cryptography*, Athens, Greece, 2010. To appear.
- [4] Jacques M. Bahi and Christophe Guyeux. A chaos-based approach for information hiding security. arXiv N° 0034939, April 2010.
- [5] J. Banks, J. Brooks, G. Cairns, and P. Stacey. On devaney’s definition of chaos. *Amer. Math. Monthly*, 99:332–334, 1992.
- [6] Christian Cachin. An information-theoretic model for steganography. *Information and Computation*, 192:41 – 56, 2004.
- [7] Francois Cayre, Caroline Fontaine, and Teddy Furon. Kerckhoffs-based embedding security classes for woa data hiding. *IEEE Transactions on Information Forensics and Security*, 3(1):1–15, 2008.
- [8] Robert L. Devaney. *An Introduction to Chaotic Dynamical Systems*. Addison-Wesley, Redwood City, CA, 2nd edition, 1989.
- [9] Robert L. Devaney. *An Introduction to Chaotic Dynamical Systems, 2nd Edition*. Westview Pr., March 2003.
- [10] Enrico Formenti. *Automates cellulaires et chaos : de la vision topologique la vision algorithmique*. PhD thesis, École Normale Supérieure de Lyon, 1998.
- [11] Nicolas Friot, Christophe Guyeux, and Jacques Bahi. Chaotic iterations for steganography - stego-security and chaos-security. In *SECRYPT'2011, Int. Conf. on Security and Cryptography*, pages ***-***, Sevilla, Spain, July 2011. To appear.
- [12] T. Furon. Security analysis, 2002. European Project IST-1999-10987 CERTI-MARK, Deliverable D.5.5.
- [13] Christophe Guyeux, Nicolas Friot, and Jacques Bahi. Chaotic iterations versus spread-spectrum: chaos and stego security. In *IIH-MSP'10, 6-th Int. Conf. on Intelligent Information Hiding and Multimedia Signal Processing*, pages 208–211, Darmstadt, Germany, October 2010.

- [14] T. Kalker. Considerations on watermarking security. In *Multimedia Signal Processing, 2001 IEEE Fourth Workshop on*, pages 201–206, 2001.
- [15] Yongyi Mao and Xiang Chen. An encryption algorithm of chaos based on sine square mapping. In *Proceedings of the 2011 Fourth International Symposium on Computational Intelligence and Design - Volume 01*, ISCID '11, pages 131–134, Washington, DC, USA, 2011. IEEE Computer Society.
- [16] J. A. Martínez-Ñonthe, A. Díaz-Méndez, M. Cruz-Irisson, L. Palacios-Luengas, J. L. Del-Río-Correa, and R. Vázquez-Medina. Cryptosystem with one dimensional chaotic maps. In *Proceedings of the 4th international conference on Computational intelligence in security for information systems*, CISIS'11, pages 190–197, Berlin, Heidelberg, 2011. Springer-Verlag.
- [17] Thomas Mittelholzer. An information-theoretic approach to steganography and watermarking. In Andreas Pfitzmann, editor, *Information Hiding*, volume 1768 of *Lecture Notes in Computer Science*, pages 1–16, Dresden, Germany, September 29 - October 1. 1999. Springer.
- [18] Luis Perez-Freire, F. Pérez-gonzalez, and Pedro Comesañ. Secret dither estimation in lattice-quantization data hiding: A set-membership approach. In Edward J. Delp and Ping W. Wong, editors, *Security, Steganography, and Watermarking of Multimedia Contents*, San Jose, California, USA, January 2006. SPIE.
- [19] Laurent Schwartz. *Analyse: topologie générale et analyse fonctionnelle*. Hermann, 1980.
- [20] Claude E. Shannon. Communication theory of secrecy systems. *Bell Systems Technical Journal*, 28:656–715, 1949.
- [21] Gustavus J. Simmons. The prisoners' problem and the subliminal channel. In *Advances in Cryptology, Proc. CRYPTO'83*, pages 51–67, 1984.
- [22] Yong Wang, Kwok-Wo Wong, Xiaofeng Liao, and Tao Xiang. A block cipher with dynamic s-boxes based on tent map. *Communications in Nonlinear Science and Numerical Simulation*, 14(7):3089 – 3099, 2009.